

La Théorie sur la “voie verte” de l’Open Access

Jocelyne JERDELET, Responsable de l’Unité Pré-tirages,
Service de l’Information Scientifique, Section Open Access,
CERN¹ 1211 Genève 23, Suisse
jocelyne.jerdelet@cern.ch

Sandrine REYES, Assistante Archiviste,
Service de l’Information Scientifique, Archives Historiques et Scientifiques,
CERN 1211 Genève 23, Suisse
sandrine.reyes@cern.ch

Résumé : Le domaine de la physique des particules a toujours œuvré pour la mise à disposition des résultats de recherche par la diffusion des pré-tirages. Ceux-ci étaient déjà distribués sous forme de document papier bien avant l’ère d’Internet.

Depuis la Convention du CERN en 1953, les pré-tirages ou prépublications rédigés par les physiciens théoriciens ont été conservés, classés puis catalogués dans la base de données du Service d’Information Scientifique (SIS) et, par la suite, numérisés. Étudiée de façon rétrospective, cette littérature grise aujourd’hui stockée sur le serveur du CERN, est mise en ligne et devient ainsi librement accessible à toute la communauté scientifique. On parle de « voie verte » de l’Open Access.

Mots-clés : Littérature grise – Pré-tirage – Prépublication – Libre accès – Dépôt institutionnel – Serveur – Voie verte

Mars 2008

¹ CERN: Organisation Européenne pour la Recherche Nucléaire <http://cern.ch>



La Théorie sur la “voie verte” de l’Open Access	1
<i>Le libre accès – Open Access (OA)</i>	<i>5</i>
<i>Open Access au CERN.....</i>	<i>5</i>
<i>Le Service d’Information Scientifique (SIS) du CERN et l’Open Access.....</i>	<i>6</i>
<i>Choix de la Série Théorie</i>	<i>7</i>
<i>Historique du catalogue de la série TH.....</i>	<i>8</i>
<i>L’objectif du projet « voie verte de l’Open Access »</i>	<i>8</i>
<i>Tâches et difficultés rencontrées.....</i>	<i>9</i>
➤ Catalogage	9
➤ Numérisation	9
➤ Standardisation et extraction de données.....	9
➤ Vérification des liens URL.....	9
➤ Recherche de Références de Publication.....	10
➤ Enquête auprès des auteurs.....	10
<i>Evaluation finale de l’analyse</i>	<i>10</i>
<i>Conclusion.....</i>	<i>11</i>
<i>Liste des serveurs de pré-tirages</i>	<i>13</i>
<i>Citations et Références</i>	<i>13</i>
<i>F I G U R E S.....</i>	<i>15</i>
➤ Fig. 1 : Production des pré-tirages par an.....	15
➤ Fig. 2 : Série Théorie par type de document.....	15
➤ Fig. 3 : Série Théorie par type de document - Graphique.....	16
➤ Fig. 4 : Résultat de la Série Théorie en Open Access.....	16
<i>A N N E X E S.....</i>	<i>17</i>
➤ Méthode de travail.....	17
➤ Procédure et tableau d’importation de nouvelles données.....	20
➤ Exemple concret :	21

Le libre accès – Open Access (OA)

Le mouvement du libre accès désigne l’ensemble des initiatives prises pour une mise à disposition des résultats de la recherche au plus grand nombre, sans restriction d’accès (libre et gratuit) pour les lecteurs, que ce soit par l’auto-archivage ou par la création de revues en libre accès.

- L’auto-archivage sur un serveur institutionnel ou sur une page personnelle, communément appelé “la voie verte”, consiste à déposer par les auteurs une copie de leur article publié dans des archives électroniques en libre accès, avec l’accord de l’éditeur.
- La publication dans des revues scientifiques en libre accès communément appelée “la voie or”, doit satisfaire à des exigences qualité, garanties par un comité de lecture. Son mode de financement permet une diffusion sans restriction d’accès et d’utilisation.

Le développement d’Internet, l’arrivée du World Wide Web (WWW), l’explosion des documents électroniques, l’augmentation du prix des abonnements des périodiques et le besoin d’accéder plus facilement et plus rapidement à l’information, sont les principaux facteurs qui ont contribué au développement de l’Open Access (OA).

Ce mouvement prend son essor en décembre 2001 à Budapest, lors d’une rencontre restreinte parrainée par l’Institut pour la Société Ouverte (Open Society Institute, OSI) [1], qui vise à harmoniser les différentes initiatives jusque là disparates.

Par la suite d’autres rencontres ont eu lieu, donnant naissance à trois textes fondateurs appelés “3B”.

« Budapest Open Access Initiative (BOAI) » [2] - Février 2002

L’initiative de Budapest recommande deux stratégies pour le libre accès :
La “voie verte” et la “voie or”.

« Bethesda Statement on Open Access Publishing (OAP) » [3] – avril 2003

La déclaration de Bethesda définit provisoirement la publication en libre accès.

« Berlin Declaration Open Access to knowledge in the Sciences and Humanities » [4] – octobre 2003

La déclaration de Berlin sur le libre accès à la connaissance en sciences exactes, en sciences de la vie, sciences humaines et sociales s’engage à promouvoir et à soutenir la diffusion gratuite de la connaissance via Internet. Cette déclaration est alors signée par le CERN.

Open Access au CERN

Depuis sa création, le CERN a toujours activement soutenu les principes de l’Open Access. Sa Convention, adoptée le 1er juillet 1953 par 12 Etats Membres², stipule que *«les résultats de ses travaux expérimentaux et théoriques [de l’Organisation] sont publiés ou de toute autre façon rendus généralement accessibles»* [5].

En 1989, Tim-Berners Lee, scientifique au CERN, invente le World Wide Web [6] qui popularise l’Internet et ouvre la voie verte de l’Open Access.

En 1991, Paul Ginsparg, physicien à Los Alamos, met en place le premier serveur de documents électroniques, nommé hep-th (High-Energy Physics – Theory), première archive en libre accès connue

² 12 Etats Membres en 1953 : Belgique, Danemark, France, Grèce, Italie, Norvège, Pays-Bas, République Fédérale Allemande, Royaume-Uni de Grande Bretagne et d’Irlande du Nord, Suède, Suisse, Yougoslavie.

aujourd’hui sous le nom d’arXiv³ ou encore le serveur de Cornell car maintenu par l’Université de Cornell.

En 1993, suivant le même exemple, le CERN ouvre son propre serveur de documents électroniques en libre accès (dépôt institutionnel) adapté aux besoins des chercheurs et des documentalistes. Les notices bibliographiques sont consultables via Internet sur ce catalogue en ligne, ainsi que le document accessible grâce à un lien hypertexte.

En 2005, le Comité de Direction du CERN adopte une nouvelle politique de libre accès concernant tous les résultats du Laboratoire, définie dans le document “Continuing CERN action on Open Access” [7]. Les auteurs sont encouragés à publier leurs travaux dans des revues scientifiques en libre accès.

En novembre 2006, les représentants des principales agences de financement européennes pour la Physique des Particules, de consortiums de bibliothèques et de la communauté scientifique se réunissent au CERN pour le lancement du projet SCOAP3⁴ (Sponsoring Consortium for Open Access Publishing in Particle Physics). Son but est de changer le mode de financement des futures publications en Physique des Particules. L’accès libre à ces futurs articles serait financé, non plus par les abonnements aux journaux souscrits annuellement par chaque bibliothèque mais par un fond commun constitué grâce au transfert de ces fonds. Déjà soutenu par plusieurs bibliothèques et agences de financement européennes, SCOAP3 s’organise maintenant à un niveau mondial.

Le Service d’Information Scientifique (SIS) du CERN et l’Open Access

Le Service d’Information Scientifique du CERN, créé en 1955, a pour mission principale d’acquérir et de gérer l’information concernant les travaux du CERN et de la diffuser à la communauté scientifique, selon la « Circulaire Opérationnelle N°6 » [8].

En 1982, le SIS informatise son catalogue par la mise en place d’un système de gestion des données bibliographiques appelé ISIS⁵ fourni par l’UNESCO⁶.

En 1989, ISIS est remplacé par le logiciel ALEPH⁷ [ExLibris, Israel]. Ce système intégré, plus adapté aux besoins de la bibliothèque du CERN, propose différents modules utiles pour traiter tous types de documents (Catalogage, Gestion des copies, Prêt, Circulation, Acquisition, ...).

L’arrivée du Web, la démocratisation d’Internet et les avancées technologiques telles que la numérisation de documents et la mise en place de serveurs de documents électroniques en libre accès ont bouleversé les milieux scientifiques et bibliothéconomiques.

Le SIS reçoit régulièrement, via des listes de diffusion, des documents papier, rédigés par des scientifiques du CERN et d’autres Instituts de recherches. Ces pré-tirages sont catalogués puis numérisés. Plus tard, la soumission électronique par l’auteur se substitue au catalogage de la bibliothèque. Très vite, le support papier disparaît au profit d’un document électronique. C’est alors que débute l’ère de la bibliothèque numérique.

³ Serveur arXiv : <http://arxiv.org/>

⁴ SCOAP3 : <http://www.scoap3.org/>

⁵ ISIS : “Integrated Set for Information” est un système de gestion de catalogage.

⁶ UNESCO : United Nations Educational, Scientific and Cultural Organization - Organisation des Nations Unies pour l’Education, la Science et la Culture.

⁷ ALEPH :” Automated Library Expandable Program” est un logiciel de gestion intégré des bibliothèques, développé par l’université de Jérusalem et produit par la société Ex-libris.

Le SIS s’adapte et développe un programme d’importation de notices bibliographiques provenant de bases de données des Instituts collaborant avec le CERN (SLAC⁸, DESY⁹...), de serveurs de pré-tirages (arXiv) et de pages Web (autres instituts) pour remplacer les listes de diffusions et pour offrir aux chercheurs un accès quasi unique à l’information. Ainsi les résultats de recherche du CERN et d’autres instituts sont mis à la disposition de la communauté scientifique sur l’interface CDS à travers de nombreuses notices bibliographiques accompagnées de leurs fichiers électroniques.

En 2006, le Service d’Information Scientifique s’investit d’avantage dans la “voie verte de l’Open Access” en mettant en ligne la série complète des documents consacrés à la physique théorique (TH) du Laboratoire.

Choix de la Série Théorie

Le choix s’est porté sur les pré-tirages de la division Théorie car cette collection est la plus ancienne (juillet 1953 à décembre 2006) et la plus prolifique du CERN, avec un total d’environ 12000 documents.

Début 2006, à la demande de la Section Gestion des Documents du SIS, ces pré-tirages sont versés par le Secrétariat de la Théorie aux Archives Historiques et Scientifiques.

Depuis 1958, cette documentation fut répertoriée manuellement par liste dans des cahiers (c.f. photo) tenus par le Secrétariat TH et identifiée par un numéro de rapport unique. Ce système de numérotation fut instauré par souci d’archivage en 1958 par Tatiana Fabergé, alors secrétaire de la division:

« Les Théoriciens publiaient beaucoup d’articles. Au début je ne réalisais pas vraiment l’importance des archives. J’attribuais quand même un numéro à chaque document. Ce fut la base du système de numérotation des publications préliminaires, toujours utilisé au CERN. » [9]



Le classement de ces documents fut produit selon une numérotation consécutive de :

CERN-TH-01 (1958) à CERN-TH-7556 (1994)

Puis de 1995 à 2003, une numérotation par année suivie d’une suite consécutive :

CERN-TH-95-001 à CERN-TH-95-352

CERN-TH-96-001 à CERN-TH-96-371

CERN-TH-97-001 à CERN-TH-97-383

CERN-TH-98-001 à CERN-TH-98-418

CERN-TH-99-001 à CERN-TH-99-415

CERN-TH-2000-001 à CERN-TH-2000-386

CERN-TH-2001-001 à CERN-TH-2001-380

CERN-TH-2002-001 à CERN-TH-2002-379

CERN-TH-2003-001 à CERN-TH-2003-316

En 2004, la structure du CERN passe de divisionnaire à départementale regroupant certaines de ses activités. La division Théorie (TH) rejoint le groupe de la Physique Expérimentale (EP) pour constituer le Département de Physique (PH). La numérotation des documents change alors pour devenir CERN-PH-TH-2004-001...

⁸ SLAC : Stanford Linear Accelerator Center, Université de Stanford, Californie, USA

⁹ DESY : Deutsches Elektronen-Synchrotron, Hamburg, Allemagne.

Historique du catalogage de la série TH

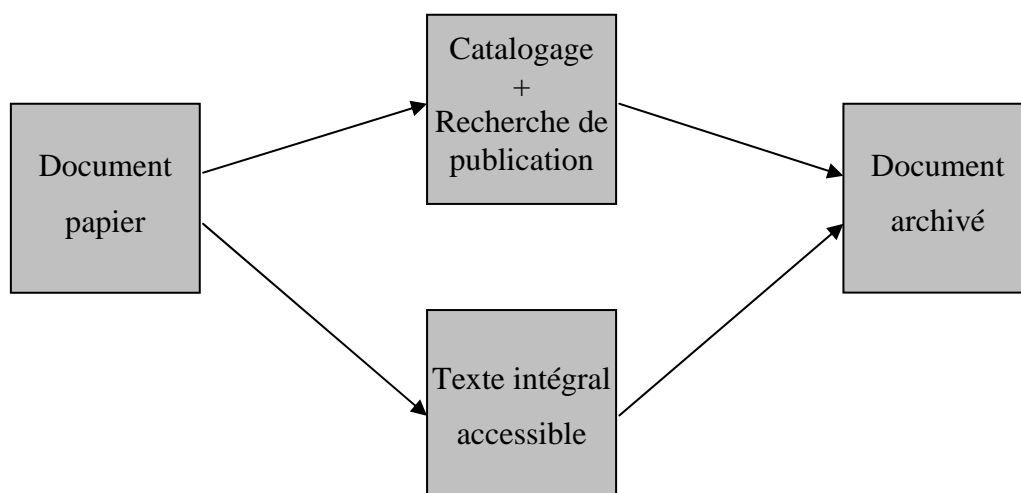
Les pré-tirages sous forme papier ne furent catalogués qu’à partir de 1982, les données sont ensuite converties en format ISBD(M)¹⁰ dans le système de gestion informatique de données ALEPH. C’est en 1990 que la numérisation commence pour les nouveaux documents.

Dès 1991, les auteurs de la théorie prennent l’habitude de soumettre leurs documents sur le serveur arXiv. A partir de 1994, pour rendre disponible cette documentation à travers le serveur du CERN, le SIS importe les métadonnées d’arXiv. Simultanément, l’institut KEK¹¹ au Japon réalise un travail considérable de numérisation de pré-tirages en physique dont certains du CERN. Le SIS obtient alors l’autorisation de KEK de lier leur numérisation à la notice bibliographique dans le catalogue en ligne du CERN.

L’objectif du projet « voie verte de l’Open Access »

Après l’analyse historique du catalogage de la collection TH, les objectifs du projet sont définis selon les critères de la voie verte de l’Open Access, et aussi selon les besoins des services du SIS et des Archives Historiques et Scientifiques :

- Pré-tirage catalogué
- Recherche de la référence de publication et mise à jour de la notice bibliographique
- Texte intégral en accès libre sur CDS
- Copie papier conservée aux Archives



¹⁰ ISBD(M) : International Standard Bibliographic Description (description bibliographique internationale normalisée). (Il s'agit d'un ensemble de normes internationales de description de catalogage définies par l'IFLA, Fédération Internationale des Associations de Bibliothécaires et d'Institutions).

¹¹ KEK : High Energy Accelerator Research Organization, Tsukuba, Japan.

Tâches et difficultés rencontrées

- Catalogage : après une première analyse de la collection, 9457 notices sont déjà répertoriées dans la base de données du SIS mais beaucoup restent à traiter, notamment les 2000 premiers documents. La perte de certaines données (telles que les numéros de rapports) suite aux diverses conversions informatiques, implique, avant tout catalogage de nouvelle notice, une recherche bibliographique par titre pour tous les numéros de rapports manquants. De plus avec l’évolution des normes bibliographiques (depuis ISBD(M) jusqu’au format MARC21¹²), l’homogénéité et la qualité du catalogage restent à vérifier sur l’ensemble des notices. Le système manuel de numérotation jusqu’en 2004 a rendu le travail du SIS plus fastidieux, car dépendant d’une recherche également manuelle.
- Numérisation : après la finalisation du catalogage des nouvelles notices, une analyse met en évidence qu’environ 50% de la collection possède un lien vers un texte intégral provenant du CERN, du serveur d’ArXiv, de KEK ou autres... Avant toute numérisation de document, une recherche dans la base de données de KEK permet une nouvelle fois l’importation de documents numérisés. Cette tâche achevée, 2527 documents restants doivent être équipés d’un code-barres, code unique d’identification sur le serveur du CERN, en vue d’être scannés.
- Standardisation et extraction de données : pour plus d’efficacité, les liens vers les fichiers des documents à numériser par le SIS sont ajoutés globalement dans les notices grâce à une extraction de données. Ainsi les numéros de rapports sont extraits de la base de données pour établir un tableau (cf. Annexes : [Tableau d’importation de nouvelles données](#)) dans lequel les codes-barres sont ajoutés pour construire le lien URL. Ces nouvelles données sont ensuite chargées dans la base.
- Vérification des liens URL : une fois la numérisation terminée, un contrôle final de tous les liens URL est effectué sur l’ensemble de la collection pour détecter d’éventuels fichiers corrompus.

Leader	001	----	-----	^nam^22^-----^a^4500
Control No. ID	003	----	----	SzGeCERN
Date and Time	005	----	----	20070607162930.0
Fixed Data	008	----	----	-----s-----r-----000^0^eng^d
Language Code	041	----	a	eng
Report No.	088	----	a	CERN-TH-01
Report No.	088	----	g	CM-P00066131
Personal Name	100	----	a	Jauch, J M
Main Title	245	----	a	Strange particle physics, pt.1
Imprint	260	----	c	1958
Imprint	269	----	a	CERN
			c	Geneva
			c	1958
Physical Des.	300	----	a	18 p
Internal Note	595	----	a	SIS:200706 Fulltext added
Internal Note	595	----	a	SIS ARCTH-2006
Subject Cat.	650	17	a	Particle Physics
			a	SzGeCERN
Local subject	690	C	a	CERN
Local subject	690	C	a	ARTICLE
A.E. Corp. Name	710	5	a	TH
Elec. Loc. Acc	856	4	u	http://preprints.cern.ch/cgi-bin/setlink?base=preprint&categ=CM-P8&id=CM-P00066131
			v	Fulltext
'Status week'	916	----	s	h
			w	200610
			y	2006
Base	BAS	----	a	13
Link	LKR	----	b	000126613
			m	book
			k	1-18
Owner	OWN	----	a	PUBLIC
Control No.	001	----	----	002601285

Catalogage de la première notice de la série Théorie au Format MARC21 pour laquelle l’objectif a été atteint :

- Les données contrôlées
- La référence de publication ajoutée
- Le document scanné
- Le lien vers le texte intégral ajouté

Notice consultable sur CDS :

<http://cdsweb.cern.ch/record/933699>

¹² MARC21 : « Machine – Readable Cataloguing » Format standard pour le stockage et l’échange de notices bibliographiques maintenu par « The Library of Congress » - Washington DC.

- Recherche de Références de Publication : afin d’enrichir le catalogue, des recherches sont entreprises dans différentes sources, bases de données ou moteurs de recherche tels que INSPEC¹³, SPIRES¹⁴, GOOGLE¹⁵ pour enrichir au maximum les notices bibliographiques avec les références de publication (journaux, comptes-rendus de conférences).
- Enquête auprès des auteurs : a la fin de l’exercice, 267 références de pré-tirages ne sont pas disponibles en Open Access car ces documents n’ont jamais été versés au Secrétariat TH. Une enquête auprès des auteurs est menée en vue d’obtenir un exemplaire papier, un fichier électronique ou une référence de publication.
La difficulté résulte essentiellement au fait que beaucoup de ces pré-tirages datent des premières années du CERN et qu’il n’est pas aisé de rentrer en contact avec ces auteurs.

Enquête auprès des auteurs		
Documents manquants		267
Adresses électroniques trouvées		167
Réponses positives avec fichiers attachés		24
Réponses positives avec copie papier		5
Réponses négatives	Pas de fichier	15
	Documents jamais finalisés	9
En attente		103

Evaluation finale de l’analyse

Depuis le début du CERN, la production de pré-tirages de la division Théorie a été en constante augmentation (cf. [Fig.1](#)). Aujourd’hui, cette collection représente 11780 notices disponibles sur CDS, parmi lesquelles :

9 457 Notices initialement existantes et vérifiées
928 Nouvelles notices cataloguées
1395 Notices recherchées et identifiées comme série TH

Une grande majorité de ces documents, soit 10145, sont des articles publiés dans des journaux scientifiques ou dans des comptes-rendus de conférences, 1579 n’ont jamais été publiés et le reste étant des documents tels que des rapports, des livres, des comptes-rendus de conférences et des thèses (cf. [Fig.2](#) et [Fig.3](#)).

2527 Documents ont été numérisés par le service du scanning du CERN.

¹³ INSPEC - <http://www.iee.org/publish/inspec/>

¹⁴ SPIRES - <http://www.slac.stanford.edu/spires/>

¹⁵ GOOGLE - <http://www.google.ch/>

Actuellement, suite aux différentes recherches, enquêtes et numérisations effectuées, 98% de la collection Théorie est disponible sur CDS en Open Access (cf. [Fig. 4](#)) :

http://cdsweb.cern.ch/search?cc=CERN-TH-Open-Access&sf=084__a&sp=TH-&so=a

L’analyse complète de la collection Théorie a nécessité 3520 heures de temps d’étude, soit 2 Equivalents Plein Temps (EPT) et 250 heures de numérisation, soit 0.15 Equivalent Plein Temps (EPT).

Conclusion

La diffusion des résultats de recherche scientifiques, d’abord effectuée par des Sociétés Savantes, est aujourd’hui diffusée, généralement, par des journaux à la publication parfois tardive et aux abonnements onéreux. Elle reste donc réservée à un ensemble de chercheurs privilégiés.

Le choix du SIS, de favoriser l’accès libre aux prépublications scientifiques, résulte d’une volonté de transmettre par un moyen efficace et économique la connaissance et l’évolution des progrès de la science. Cette mise à disposition gratuite et plusieurs mois avant la publication de ces pré-tirages, représente un intérêt particulier, notamment en Physique des Particules.

Commencé en 2006, ce projet de mise en ligne dans CDS et de conservation des documents aux Archives Historiques et Scientifiques du CERN doit s’appliquer à la totalité des collections de pré-tirages du CERN (rétrospectives et courantes). Ainsi ces prépublications pourront être réutilisées, copiées, téléchargées, imprimées, archivées, consultées à distance, offrant aux utilisateurs rapidité, fiabilité, facilité d’utilisation et gratuité de l’information.

Le développement et la généralisation de la voie verte de l’Open Access au sein des dépôts institutionnels permettraient d’assurer, préserver et pérenniser l’accès à la totalité des résultats scientifiques, sans limitation d’utilisation et sous l’unique contrôle des instituts, garantissant une efficacité, une ouverture et une simplicité dans un esprit d’économie non négligeable.

Liste des serveurs de pré-tirages

CDS : CERN Document Serveur : <http://cdsweb.cern.ch/>

ArXiv : Serveur de pré-tirages en physique des Hautes Energies, astrophysique, etc.
<http://xxx.lanl.gov/>

SLAC SPIRES : <http://www.slac.stanford.edu/spires/hep/>

DESY SPIRES : <http://www-library.desy.de/spires/hep/>

CCSD : Centre pour la Communication Scientifique Directe : <http://ccsd.cnrs.fr/>

KISS for preprints : KEK Information Service System : http://www-lib.kek.jp/KISS/kiss_preprint.html

HAL (Hyper Article en Ligne) : <http://hal.ccsd.cnrs.fr/>

Citations et Références

[1] Open Society Institute (OSI) & Soros Foundation Network
<http://www.soros.org/about>

[2] Budapest Open Access Initiative (BOAI), février 2002 -
<http://www.soros.org/openaccess/fr/read.shtml>

[3] Bethesda Statement on Open Access Publishing (OAP), avril 2003
<http://www.earlham.edu/~peters/fos/bethesda.htm>

[4] Berlin Declaration Open Access to knowledge in the Sciences and Humanities,
<http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>

[5] La Convention pour l'établissement d'une Organisation Européenne pour la Recherche Nucléaire (CERN), adoptée le 1er juillet 1953 par 12 Etats Membres
<http://cdsweb.cern.ch/record/330625>

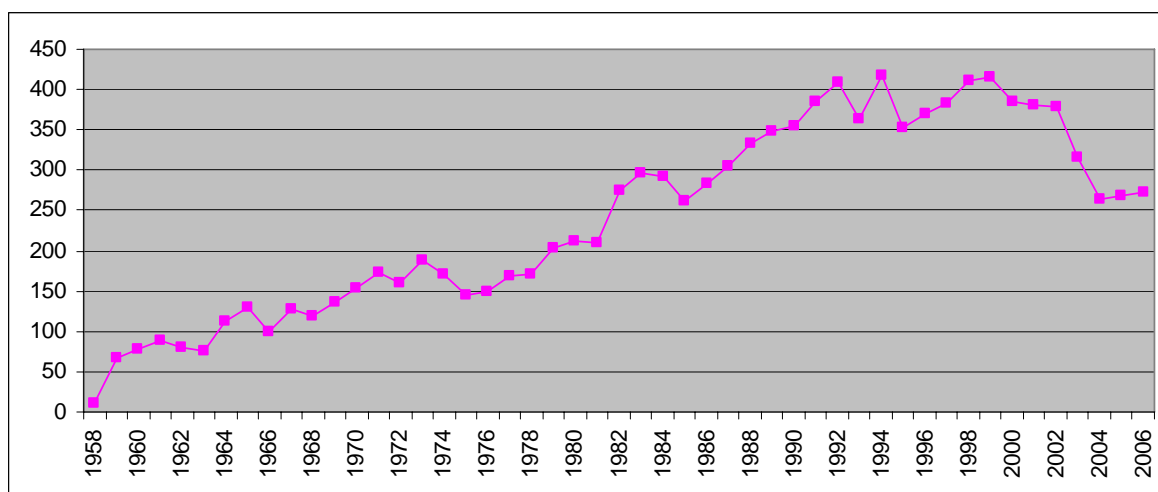
[6] Histoire du World Wide Web
<http://info.cern.ch/default-fr.html>

[7] “Continuing CERN action on Open Access” CERN-OPEN-2005-006 – Altarelli, Guido
<http://cdsweb.cern.ch/record/828991>

[8] Circulaire Opérationnelle N°6
<http://open-access.web.cern.ch/Open-Access/oc-06.pdf>

[9] Infiniment CERN, Témoins de cinquante ans de recherches, « La Théorie en Pratique » - page 21. Citation de Tatiana Fabergé, Secrétaire de la Division TH de 1957 à 1995.
<http://cdsweb.cern.ch/record/796102>

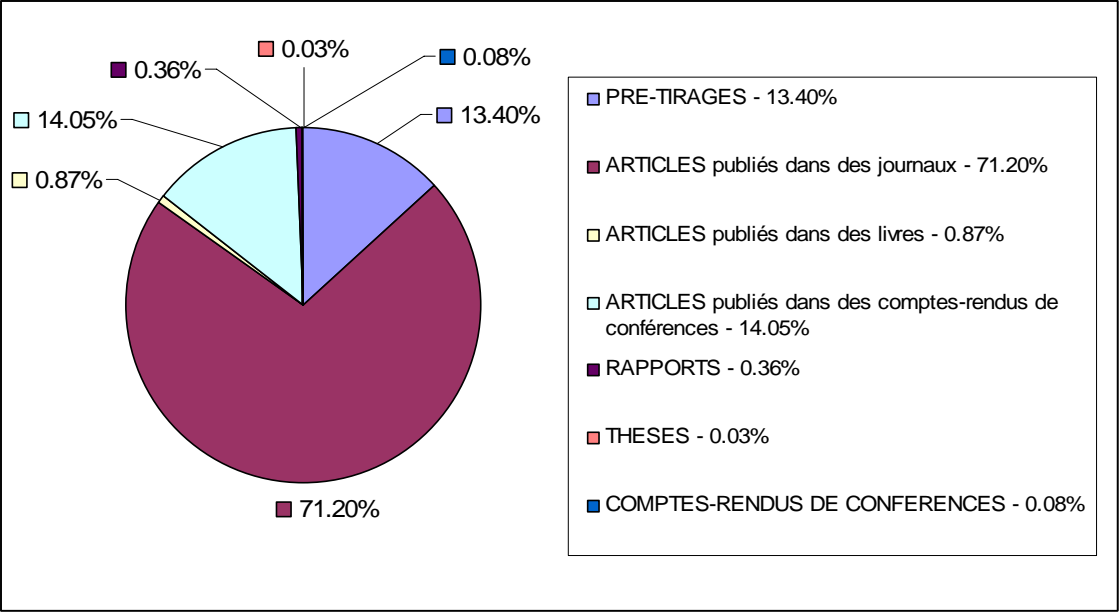
FIGURES



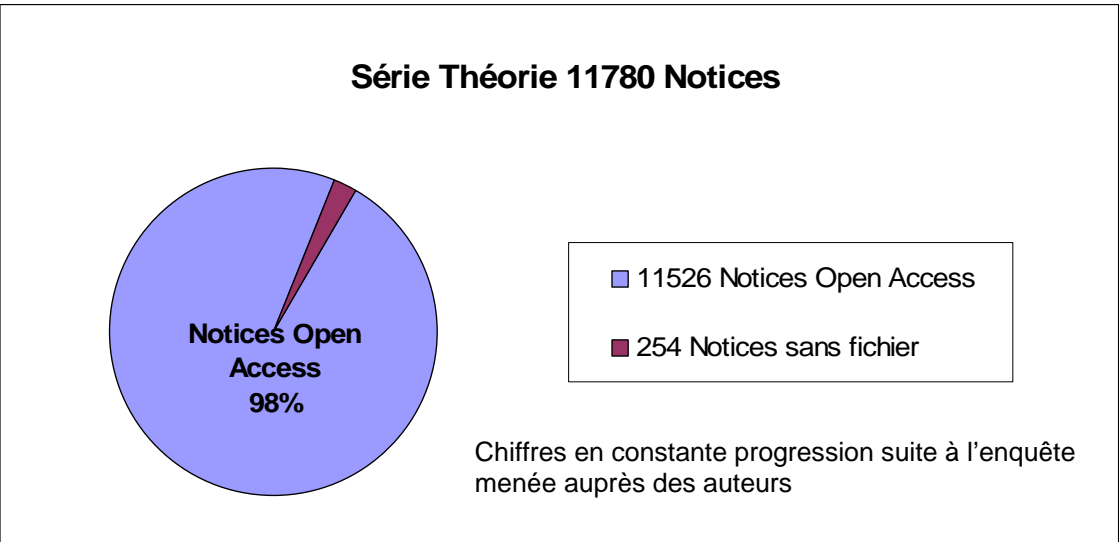
➤ [Fig. 1 : Production des pré-tirages par an](#)

Pré-tirages	1579
Articles publiés	10145
Thèses	4
Rapports - Livres	42
Comptes-rendus de conférences publiés dans un livre	9
Comptes-rendus de conférences publiés dans un journal	1
Notices consultables dans CDS	11780

➤ [Fig. 2 : Série Théorie par type de document](#)



➤ [Fig. 3 : Série Théorie par type de document - Graphique](#)



➤ [Fig. 4 : Résultat de la Série Théorie en Open Access](#)

ANNEXES

➤ Méthode de travail

D'après une première étude de la série nous avons pu constater différents cas de figures :

Cas de figures	Support	Notice bibliographique	Texte intégral ou URL
1	OUI	OUI	OUI
2	OUI	OUI	NON
3aI	OUI	NON	OUI
3aII & 3b	OUI	NON	NON
4a	NON	OUI	OUI
4b	NON	OUI	NON
5aI	NON	NON	OUI
5aII & 5b	NON	NON	NON

1. Nous avons une copie papier, une notice bibliographique et un texte intégral ou URL existant. L'objectif est atteint et le document est donc archivé comme copie de conservation.
2. Nous avons une copie papier, une notice bibliographique mais pas de texte intégral. Un code-barres est ajouté au document pour la numérisation et l'URL ajoutée dans la notice
3. Nous n'avons qu'une copie papier. Une recherche par titre est effectuée dans la base :
 - a. Si la notice est trouvée, nous ajoutons le champ correspondant à la numérotation de l'article "CERN-TH-*". nous procédons ensuite à la vérification éventuelle du texte intégral.
 - I. Si nous avons une URL, le document est alors archivé comme copie de conservation
 - II. Si nous n'avons pas d'URL, le document est équipé d'un code-barres pour la numérisation et l'URL est ajoutée à la notice.
 - b. Si le titre est introuvable, le document est alors à cataloguer, à équiper d'un code-barres et à numériser.
4. Nous n'avons pas de copie papier mais une notice bibliographique dans la base :
 - a. Si cette notice bibliographique possède une adresse URL alors le document est imprimé grâce à la version électronique et inséré dans la série.
 - b. Si cette notice bibliographique ne possède pas d'adresse URL, une note est ajoutée dans la notice pour informer le lecteur que le document est manquant. "*Not Held by the CERN Library*" et la notice est imprimée pour l'insérer dans la série. Une demande est envoyée à l'auteur en vue d'obtenir un pré-tirage papier ou électronique ainsi qu'une référence de publication.

5. Nous n'avons aucune copie papier, aucune notice bibliographique, aucun texte intégral. Nous consultons les listes de diffusions (cahier du secrétariat) en nous référant au numéro de rapport pour obtenir un titre.
 - a. Si la référence est trouvée dans les cahiers, nous faisons une recherche par titre et par auteur dans la base :
 - I. Si la notice est trouvée, nous ajoutons le champ correspondant à la numérotation de l'article "CERN-TH-*", nous procédons ensuite à la vérification du texte intégral afin de le numériser si nécessaire en ayant préalablement ajouté le code-barres et mis à jour la notice (code-barres + URL). Une demande est envoyée à l'auteur en vue d'obtenir un pré-tirage papier ou électronique ainsi qu'une référence de publication.
 - II. Si le titre est introuvable dans la base, le document est à cataloguer de manière succincte, d'après les informations du cahier en ajoutant une note pour informer le lecteur que le document est manquant, "*Not held by the CERN Library*". Une demande est envoyée à l'auteur en vue d'obtenir un pré-tirage papier ou électronique ainsi qu'une référence de publication.
 - b. Si le titre est introuvable dans les cahiers, nous cataloguons une numérotation et non pas un document dans une base à part non-accessible sur le catalogue en ligne. Cette procédure a été mise en place pour obtenir une liste complète de numéros de rapports et pour éviter d'analyser une seconde fois ces mêmes références.

	Copie papier	Notice	URL ou Texte intégral	Action				Mise à jour de la notice	Texte Intégral	Copie papier	Résultat	
1	OUI	OUI	OUI						Vérification du texte intégral		Conservation du <u>document</u> aux archives	
2	OUI	OUI	NON	Obtenir un fichier électronique				Champs du code-barres et de l'adresse URL à rajouter	Numérisation	Equiper d'un code-barres	Conservation du <u>document</u> aux archives	
3 aI	OUI	NON		Recherche par titre dans la base	OUI	Avec URL		Ajout du champ correspondant au numéro de rapport	Vérification du texte intégral		Conservation du <u>document</u> aux archives	
3 aII						Sans URL		Ajout du champ correspondant au numéro de rapport + code-barres + URL	Numérisation	Equiper d'un code-barres	Conservation du <u>document</u> aux archives	
3 b					NON		Notice à cataloguer avec code-barres + URL	Numérisation	Equiper d'un code-barres	Conservation du <u>document</u> aux archives		
4 a	NON	OUI	OUI	Obtenir une copie papier					Vérification du texte intégral	Imprimer le <u>document</u> pour l'insérer dans la série	Conservation du <u>document</u> aux archives	
4 b			NON	Obtenir une copie papier et un fichier électronique				Ajouter la note “ <i>Not held by the CERN Library</i> ”		Imprimer la <u>notice</u> pour l'insérer dans la série	Enquête auprès des auteurs	
5 aI	NON	NON		Recherche par numéro de rapport le titre dans les cahiers	OUI	Recherche par titre dans la base	OUI	Avec URL	Ajout du champ correspondant au numéro de rapport	Vérification du texte intégral	Imprimer le <u>document</u> pour l'insérer dans la série	Conservation du <u>document</u> aux archives
								Sans URL	Ajout du champ correspondant au numéro de rapport et ajouter la note « <i>Not held by the CERN Library</i> »		Imprimer la <u>notice</u> pour l'insérer dans la série	Enquête auprès des auteurs
5 aII							NON		Notice à cataloguer avec le numéro de rapport + titre et Ajouter la note “ <i>Not held by the CERN Library</i> ”			
5 b					NON				Enregistrer le numéro de rapport dans une base à part et Ajouter la note “ <i>Not held by the CERN Library</i> ”		Imprimer la <u>notice</u> pour l'insérer dans la série	Conservation de la <u>notice</u> aux archives

➤ **Procédure et tableau d'importation de nouvelles données**

Le procédé consiste à travailler boîte par boîte pour :

1. inventorier les copies papier et les notices bibliographiques par ordre de numéro de rapport et de façon consécutive,
2. extraire de la base de données les numéros de rapports et les URL des documents concernés,
3. sortir des boîtes les documents qui n'ont pas d'URL pour les équiper de code-barres,
4. créer un tableau sur Microsoft Office Word avec ces informations,
5. utiliser le lecteur de code-barres, pour entrer les données dans la colonne 6 du tableau ci-dessous,
6. copier la colonne 6 sur la colonne 10 qui servira aux URL des documents,
7. créer un nouveau fichier avec ces données pour le soumettre au système, ceci dans le but d'ajouter globalement les champs concernant l'URL et le numéro de code-barres aux notices (fichier Excel transformé en fichier texte sous UNIX – voir tableaux ci-dessous),
8. une fois la numérisation terminée, reclasser les documents dans les boîtes,
9. réunir toutes les notices contenant la note « Not held by the CERN Library » pour demander aux auteurs une copie des documents manquants.

1	2	3	4	5	6	7	8	9	10
Numéro de système	Numéro du champ pour la référence du rapport	CERN-TH-XX	Numéro de système	Numéro du champ pour le Code-barres	Code-barres	Numéro de système	Numéro du champ pour l'adresse URL	http://preprints.cern.ch/cgi-bin/setlink?base=preprint&categ=CM-P&id=	Code-barres
Numéro de système	Numéro du champ pour la référence du rapport	CERN-TH-XX	Numéro de système	Numéro du champ pour le Code-barres	Code-barres	Numéro de système	Numéro du champ pour l'adresse URL	http://preprints.cern.ch/cgi-bin/setlink?base=preprint&categ=CM-P&id=	Code-barres
Numéro de système	Numéro du champ pour la référence du rapport	CERN-TH-XX	Numéro de système	Numéro du champ pour le Code-barres	Code-barres	Numéro de système	Numéro du champ pour l'adresse URL	http://preprints.cern.ch/cgi-bin/setlink?base=preprint&categ=CM-P&id=	Code-barres

Tableau d'importation de nouvelles données

➤ Exemple concret :

Tableau concernant les documents à numériser

1	2	3	4	5	6	7	8	9	10
000045260	088 L \$\$a	CERN-TH-30	000045260	088 L \$\$9	CM-P00058610	000045260	8564 L \$\$u	http://preprints.cern.ch/cgi-bin/setlink?base=preprint&categ=CM-P&id=	CM-P00058610
000045261	088 L \$\$a	CERN-TH-31	000045261	088 L \$\$9	CM-P00058611	000045261	8564 L \$\$u	http://preprints.cern.ch/cgi-bin/setlink?base=preprint&categ=CM-P&id=	CM-P00058611
000045262	088 L \$\$a	CERN-TH-32	000045262	088 L \$\$9	CM-P00058612	000045262	8564 L \$\$u	http://preprints.cern.ch/cgi-bin/setlink?base=preprint&categ=CM-P&id=	CM-P00058612

1^{er} Fichier emacs sous Linux créé avec les données concernant la référence du code-barres et lien URL.
Les 3 premières lignes représentent le contenu des colonnes N° 4, 5, 6 du tableau ci-dessus
Les 3 lignes suivantes représentent le contenu des colonnes N°7, 8, 9 et 10 du tableau ci-dessus

2^{ème} fichier emacs (Linux-UNIX) trié par ordre de numéro de système pour intégrer les données à la base.